Abhinav Maheshwari

abhinavm@umd.edu | +1 (240) 501-3592 | LinkedIn | GitHub

EDUCATION

Master of Science in Data Science and Analytics — **Aug 2023 – May 2025**
University of Maryland — *College Park, MD*
- Graduate Teaching Assistant – INST 447, Data Sources and Manipulation

Bachelor of Engineering in Electrical and Electronics — **Aug 2017 – May 2021**
Birla Institute of Technology — *Ranchi, India*

TECHNICAL SKILLS

Programming Languages: Python, R, C, MATLAB, SQL
Frameworks & Libraries: TensorFlow, PyTorch, Keras, Scikit-learn, LangChain, NumPy, Pandas, SciPy, NLTK, HuggingFace, RLHF
Big Data & Cloud Technologies: Apache Kafka, Google Cloud Platform (GCP), Amazon Web Services (AWS)
Data Visualization: Tableau, Power BI, Matplotlib, Seaborn
Core Expertise: Credit Risk Modeling, Predictive Analytics, Fraud Detection, Machine Learning, NLP, Experimentation & A/B Testing
Development Tools: Git, Linux, Docker, Jupyter

PROFESSIONAL EXPERIENCE

AI Engineer Intern — **Jun 2024 – Aug 2024**
Reality AI Lab — *New York, NY*
- Designed and deployed an LLM-powered RAG system using LangChain and transformer models, boosting context-aware retrieval accuracy by 50% in production settings.
- Built agentic workflows leveraging real-time document ingestion and multi-step prompt chains, reducing redundant manual queries and increasing system efficiency by 60%.
- Deployed containerized inference APIs via Google Cloud's Vertex AI, achieving 40% latency reduction and enabling production readiness for generative models.

Data Scientist — **Aug 2021 – Jul 2023**
Venkateshwar Ispat — *Raipur, India*
- Built real-time data pipelines using Python and SQL that processed production telemetry and behavioral logs from 20+ sensors, enabling predictive maintenance and reducing downtime by 25%.
- Deployed scikit-learn models for anomaly detection and optimization, increasing daily steel throughput by 20 metric tons.
- Led a KPI tracking initiative by integrating Power BI dashboards with live SQL feeds, delivering product performance insights that accelerated leadership decision-making by 20%.

Data Science Intern — **May 2020 – Jul 2020**
Corizo — *Remote, India*
- Built modular data preprocessing pipelines in Python to clean and structure diverse datasets, reducing model training errors by 15% and improving baseline accuracy by 10%.
- Applied linear regression, random forest, and gradient boosting models to uncover business insights, increasing client decision-making efficiency by 20%.
- Conducted exploratory data analysis (EDA) to identify key behavioral trends and patterns, informing early-stage product experiments.

PROJECTS

RetentionCloud
- Architected an AWS-hosted ML pipeline for churn prediction and customer lifetime value (CLV) forecasting using XGBoost and SMOTE, improving retention strategy precision by 30%.
- Integrated Lambda + SageMaker for low-latency, event-driven inference, serving real-time predictions to a Streamlit dashboard used by business teams.
- Designed and deployed user segmentation models that identified high-risk, high-value customers, increasing targeted retention campaign ROI by 22%.

Clinical Trial Sentiment Analysis
- Designed a Python-powered sentiment analysis system to extract insights from Reddit comments on clinical trials, utilizing PRAW for API interaction, NLTK's VADER for NLP, and OpenAI's GPT-3.5 for AI-generated responses.
- Processed large-scale unstructured social media data from Reddit, identifying sentiment patterns and ethical considerations in medical discussions.
- Built automated pipelines to continuously ingest and analyze social discourse, enabling longitudinal studies on perception shifts and content impact.

AI-Driven Quiz Generation
- Engineered an NLP-powered quiz generation system using Google Cloud's Vertex AI, leveraging Gemini Pro for context-aware question formulation.
- Implemented a robust document processing and embedding pipeline with LangChain and ChromaDB, facilitating efficient storage and retrieval for question generation.
- Developed an interactive Streamlit-based interface for seamless PDF ingestion, topic selection, and quiz navigation.